



## Cost-Effective Design of Multiresponse Experiments

Sergei LEONOV

(Research Statistics Unit, Collegeville, PA)

Joint work with Valerii FEDOROV

NIST Conference on Designs for GLMs

April 18-20, 2002

### Outline

- Practical problems (rationale)
- Statistical problem
- Parameter estimation
- Optimal design, equivalence theorem
- Examples, standard normalization
- Optimal design with cost constraints, examples

### Keywords

Multiple responses, Variance depending on unknown parameters,  
Cost constraints

3

4

### Rationale: Biopharmaceutical Applications

#### Motivating Example 1: Quantal models for dose response

$$Y = Y(x) = \begin{cases} 1, & \text{a response is present on dose } x, \\ 0, & \text{no response,} \end{cases}$$

Probability of response:  $P\{Y = 1|x\} = \eta(x, \theta)$ ,

$\eta(x, \theta) = \pi(\theta_1 + \theta_2 x)$ ,  $\pi$  - p.d.f. (logit, probit, etc.).

D-optimality, c-optimality (estimation of percentiles  $ED_p$ ):  
optimal designs are invariant in the induced space  $z = \theta_1 + \theta_2 x$ , but  
depend on  $\theta$  in the original  $x$ -space.

White (1975), Wu (1988), Ford, Torsney, and Wu (1992)

↓

GLM Framework

#### Example 2: Multiparameter logistic models for bioassays

$$E(y|x; \theta) = \eta(x, \theta) = \theta_1 + \frac{\theta_2 - \theta_1}{1 + \exp(\theta_3 + \theta_4 x)},$$

If  $\theta_1 = A$ ,  $\theta_2 = B$  (Min and Max of observations), then

$$\theta_3 + \theta_4 x = \log \frac{B - \eta}{\eta - A}$$

↓

reduced to GLM Framework

Power model of variance:  $\text{Var}(y|x, \theta) = S(x, \theta) = \delta_1 \eta^{\delta_2}(x, \theta)$ .

Karpinski (1990), Hedayat et al (1997)

Motivating Example 3: PK model used in one of GSK studies

Measuring drug concentration at  $k$  different time points:

$$\eta(x_1, \theta), \eta(x_2, \theta), \dots, \eta(x_k, \theta),$$

$x_i$  - time of  $i$ -th sample;

$\theta$  - PK parameters,  $\theta \sim \mathcal{F}(\theta)$  - population model (random coefficients, or mixed effects model).

Costs depend on the number of samples ( $k = 2$ ):

$c_1(x)$  - cost of single sample at time  $x$ ,

$c_2(x_1, x_2)$  - cost of two samples at  $x_1, x_2$ .

$k = 2$ : the options are

- Take one sample:  $\eta_1(x, \theta) + \text{error} \Leftrightarrow \text{cost } c_1(x)$ ,
  - Take two samples:  $\eta_1, \eta_2 + \text{error} \Leftrightarrow \text{cost } c_2(x)$ ,
- plus take into account the covariance structure

$$\Downarrow$$

$$E[y|x, \theta] = \eta(x, \theta), \quad \text{Var}[y|x, \theta] = S(x, \theta).$$

Questions: (1) How many samples to take: one or two?

(2) At which times  $\{x\}$ ?

PK Model for oral drug: two-compartment model with first-order absorption and elimination from central compartment,

$$\begin{cases} \dot{f}_0(x) = -K_a f_0(x) \\ \dot{f}_1(x) = K_a f_0(x) - (K_{12} + K_{10}) f_1(x) + K_{21} f_2(x) \\ \dot{f}_2(x) = K_{12} f_1(x) - K_{21} f_2(x) \end{cases}$$

Initial condition  $f(0) = (D, 0, 0) \implies$

$$\boxed{\eta(x, \theta) = f_1(x)/V_1}$$

$V_1$  - volume of distribution,  $K_a$  - absorption rate constant,  
 $K_{10}, K_{12}, K_{21}$  - transfer rate constants.



Nonlinear model with multiple responses

All Examples: variance depends on unknown parameters

Regression Model

$\{y_i, i = 1, \dots, N\}$  - observations,

$$E[y|x] = \eta(x, \theta), \quad \text{Var}[y|x] = S(x, \theta),$$

$\eta(x, \theta)$  is  $(k \times 1)$  vector;  $S(x, \theta)$  is  $(k \times k)$  matrix.

$x \in \mathcal{X} \subset R^l$ : independent variables (predictors),

$\theta \in \Omega \subset R^m$ : m unknown parameters.

In Example 3:  $k = 1$  or  $2$ ;  $l = 1$  (time).

**Information Matrix**  $\mu(x, \theta)$ :

Gaussian  $y_i$ , single response,  $k = 1$ :

$$\mu(x, \theta) = \frac{1}{S} \frac{\partial \eta}{\partial \theta} \frac{\partial \eta}{\partial \theta^T} + \frac{1}{2S^2} \frac{\partial S}{\partial \theta} \frac{\partial S}{\partial \theta^T}.$$

$$S = S(x, \theta), \quad \eta = \eta(x, \theta)$$

Multiple Responses,  $k > 1$ :  $\mu(x, \theta) = [\mu_{\alpha\beta}(x, \theta)]_{\alpha, \beta=1}^m$ ,

$$\mu_{\alpha\beta}(x, \theta) = \frac{\partial \eta}{\partial \theta_\alpha} S^{-1} \frac{\partial \eta}{\partial \theta_\beta} + \frac{1}{2} \text{tr} \left[ S^{-1} \frac{\partial S}{\partial \theta_\alpha} S^{-1} \frac{\partial S}{\partial \theta_\beta} \right].$$

Magnus and Neudecker (1988)

Maximum likelihood estimator (MLE):

$$\theta_N = \arg \min_{\theta \in \Omega} \sum_{i=1}^N L(y_i | x_i, \theta),$$

$$L(y_i | x, \theta) = \log |S(x_i, \theta)| + [y_i - \eta(x_i, \theta)]^T S^{-1}(x_i, \theta) [y_i - \eta(x_i, \theta)].$$

Convergence of the MLE: regularity conditions  $\implies$

$$\sqrt{N}(\theta_N - \theta^*) \sim \mathcal{N}(0, M^{-1}(\theta^*)), \quad M(\theta^*) = \lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N \mu(x_i, \theta^*)$$

Nonlinear least squares: Jennrich (1969), Wu (1981)

### Iterated Estimators

(a)  $S = S(x_i) \implies$  well studied Generalized LS,

$$\hat{\theta}_{1N} = \arg \min_{\theta} \sum_{i=1}^N \frac{[y_i - \eta(x_i, \theta)]^2}{S(x_i)}. \quad (E1)$$

(b) When  $S = S(x, \theta)$ , it seems natural to replace (E1) by

$$\tilde{\theta}_{2N} = \arg \min_{\theta} \sum_{i=1}^N \frac{[y_i - \eta(x_i, \theta)]^2}{S(x_i, \theta)}, \quad (E2)$$

but  $\tilde{\theta}_N$  in general is not consistent.

Fedorov (1974), Beal & Sheiner (1988), Malyutov (1988), Vonesh & Chinchilli (1997)

(c) Iteratively reweighted least squares (IRLS)

$$\tilde{\theta}_{3N} = \lim_{N \rightarrow \infty} \theta_t, \quad (E3)$$

$$\theta_t = \arg \min_{\theta} \sum_{i=1}^N [y_i - \eta(x_i, \theta)]^T S^{-1}(x_i, \theta_{t-1}) [y_i - \eta(x_i, \theta)]$$

$\tilde{\theta}_{3N}$  is strongly consistent,

$$\text{Var}[\tilde{\theta}_{3N}] \simeq \left[ \sum_{i=1}^N \frac{\partial \eta(x_i, \theta)}{\partial \theta} S^{-1}(x_i, \theta) \frac{\partial \eta(x_i, \theta)}{\partial \theta^T} \right]^{-1} \Big|_{\theta=\tilde{\theta}_{3N}} \implies$$

inferior to the MLE.

### Combined Iteratively Reweighted LS estimator (CIRLS)

Includes 'squared' deviations of the predicted variance matrices  $S(x, \theta)$  from observed residual matrices:

$$\hat{\theta}_N = \lim_{t \rightarrow \infty} \theta_t, \quad \theta_t = \arg \min_{\theta} R_N(\theta, \theta_{t-1}), \quad (E4)$$

$$R_N(\theta, \theta') = \sum_{i=1}^N [y_i - \eta(x_i, \theta)]^T S^{-1}(x_i, \theta') [y_i - \eta(x_i, \theta)] + \\ + \frac{1}{2} \sum_{i=1}^N \text{tr} \left[ \{ [y_i - \eta(x_i, \theta')] [y_i - \eta(x_i, \theta')]^T - \Delta \eta(x_i; \theta, \theta') - S(x_i, \theta) \} S^{-1}(x_i, \theta') \}^2 \right],$$

where

$$\Delta \eta(x_i; \theta, \theta') = [\eta(x_i, \theta) - \eta(x_i, \theta')] [\eta(x_i, \theta) - \eta(x_i, \theta')]^T.$$

The MLE  $\theta_N$  and CIRLS  $\hat{\theta}_N$  are asymptotically equivalent

Fedorov, Gagnon, and Leonov (2001)

### Optimal Design

$n_i$  independent measurements at  $x_i$ ,  $\sum_{i=1}^n n_i = N \implies$

Normalized information matrix:

$$M_\xi(\theta) = \frac{1}{N} \sum_{i=1}^n n_i \mu(x_i, \theta) = \sum_{i=1}^n w_i \mu(x_i, \theta)$$

$\xi = \{w_i, x_i\}$  - design;  $w_i = n_i/N$ .

Asymptotic variance of the MLE:  $\text{Var}(\hat{\theta}) \approx M_\xi^{-1}(\hat{\theta})/N$ .

Thus, 'minimize' variance:  $\xi^* = \arg \min_{\xi} \Psi[M_\xi^{-1}(\theta)]$ ,

subject to  $\{0 \leq w_i \leq 1, \sum_{i=1}^n w_i = 1; x_i \in \mathcal{X}$  - design region $\}$ .

How to get frequencies:  $n_i \approx N w_i$ .

- D-criterion:  $\Psi = \log |M_\xi^{-1}(\theta)|$  ( $\sim V$ (confidence ellipsoid))

- A-criterion:  $\Psi = \text{tr}[AM_\xi^{-1}(\theta)]$  (eg., average variance of  $\hat{\theta}_i$ )

### Generalized Equivalence Theorem

D-optimality: take measurements at points where the normalized variance of prediction (sensitivity function) is the worst.

$$(1) \quad x_i \Leftarrow \max_x \text{tr}[\mu(x, \theta) M_\xi^{-1}(\theta)] \quad (\text{algorithm}).$$

$$(2) \quad \psi(x, \theta, \xi^*) = \text{tr}[\mu(x, \theta) M_{\xi^*}^{-1}(\theta)] \leq m, \quad m = \dim(\theta).$$

$\xi^*$  - D-optimal design;  $\psi(x, \theta, \xi)$  - sensitivity function.

A-optimality:  $\psi(x, \xi, \theta) = \text{tr}[\mu(x, \theta) M_\xi^{-1}(\theta) A M_\xi^{-1}(\theta)]$

### D-optimal designs with a single response

1. Was  $S(x)$ :

$$x_i \Leftarrow \max_x \frac{\text{Var}[\hat{\eta}(x, \theta)]}{S(x)}.$$

2. Now  $S(x, \theta)$ :

$$x_i \Leftarrow \max_x \left\{ \frac{\text{Var}[\hat{\eta}(x, \theta)]}{S(x, \theta)} + \frac{\text{Var}[\hat{S}(x, \theta)]}{2S^2(x, \theta)} \right\} = \\ = \frac{1}{S(x, \theta)} \frac{\partial \eta(x, \theta)}{\partial \theta^T} M_\xi^{-1}(\theta) \frac{\partial \eta(x, \theta)}{\partial \theta} + \frac{1}{2S^2(x, \theta)} \frac{\partial S(x, \theta)}{\partial \theta^T} M_\xi^{-1}(\theta) \frac{\partial S(x, \theta)}{\partial \theta}.$$

Atkinson and Cook (1995), Downing, Fedorov, Leonov (2001)

Often the “information matrix” contains the first term only  $\Rightarrow$   
 corresponds to the asymptotic covariance matrix of IRLS -  
 not covariance matrix of the MLE!

Bezeau and Endrenyi (1986)

Special case: variance is proportional to the square of the mean,  
 $S(x, \theta) = a\eta^2(x, \theta)$ ,  $a$  - given constant.

For a single Gaussian observation,

$$\mu(x, \theta) = \frac{f(x, \theta) f^T(x, \theta)}{\eta^2(x, \theta)} \left[ \frac{1}{a} + 2 \right], \quad f(x, \theta) = \frac{\partial \eta(x, \theta)}{\partial \theta}.$$

So, the two matrices differ by the constant multiplier  $\Rightarrow$   
 optimal designs are the same.

## Locally Optimal Designs

Optimal design  $\xi^*$  depends on:

- Model {response  $\eta(x, \theta)$  and variance  $S(x, \theta)$ }
- Design region  $\mathcal{X}$
- Parameter estimates  $\theta$  for nonlinear models  $\Rightarrow$  locally optimal designs

Alternative approaches: Bayesian, minimax, adaptive

Nevertheless, locally optimal designs provide a reference point for other candidate designs.

**Example S1, four-parameter logistic model:** how a compound inhibits the proliferation of “bad” cells in a cell-based assay.

response:  $\eta(x, \theta) = \theta_1 + \frac{\theta_2 - \theta_1}{1 + (x/\theta_3)^{\theta_4}}$  (cf. p.4)

variance:  $S(x, \theta) = \theta_5 \eta^{\theta_6}(x, \theta)$  (power model)

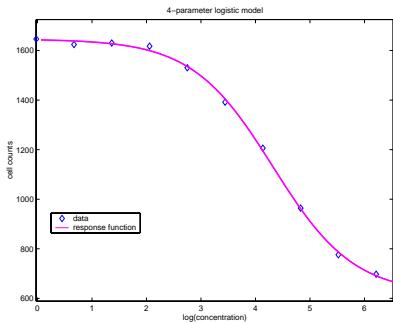


Figure 1: 4-parameter logistic model, two-fold dilutions

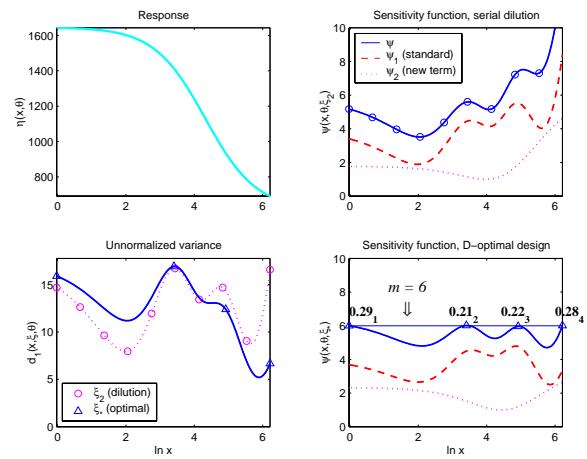


Figure 2: Sensitivity function: 6 parameters  $\Rightarrow$  4 points.  $\text{Eff}_{\xi_2} = 0.83$

Design's efficiency :  $\text{Eff}_\xi = [|M_\xi| / |M_{\xi^*}|]^{1/m}$ .

Alternative variance model:

$$S(x, \theta) = \theta_5 + \theta_6 \eta(x, \theta)$$

Questions:

- (1) How far is 2-fold dilution from the optimal design?
- (2) Are there other “good” serial dilutions  $\xi_a$ ?

When dilution design does not cover the whole range of the curve, its efficiency drops dramatically (over 40%)

Limitations of locally optimal designs:

- Relative ease in making serial dilutions in 96-well plate assay,
- Extra time for making the appropriate D-optimal design



Thus, compared dilution designs with different dilution factors  $\Rightarrow$  switched to 2.5 (3)-fold dilutions (covering wider range of concentrations)

**Example S2: measuring fluorescence curves**

(thrombin generation)

Two different models can be used:

Exponential

$$\eta = \theta_1 + \theta_2 \exp(\theta_3 t^{\theta_4})$$

Skewed logistic

$$\eta = \theta_1 + \frac{y_{min} - \theta_1}{[1 + \exp(\theta_2 + \theta_3 t)]^{\theta_4}}$$

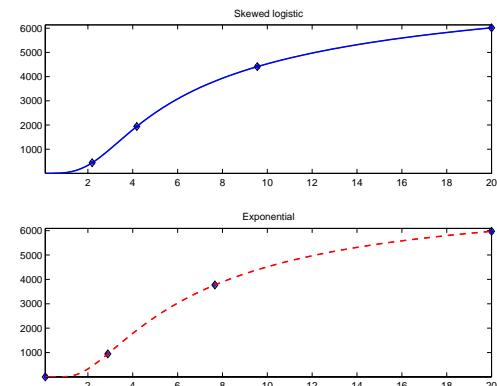


Figure 3: Two models

Equivalence theorem: design region  $X=[0,20]$

Traditional design - sampling every min ( $w_i=1/21$ ).

Locally optimal designs: 4 points ( $w_i=1/4$ ) for both models, but

- (1) depend on  $\theta$ ; (2) different for 2 models

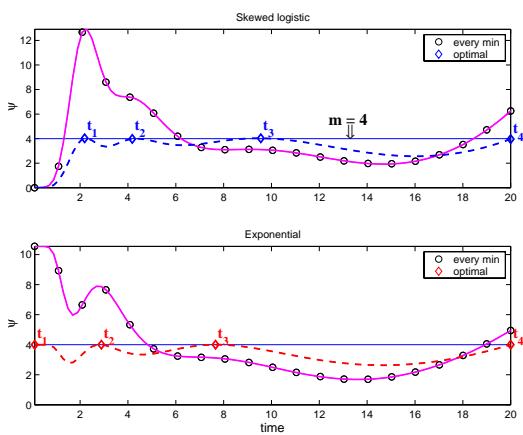


Figure 4: Sampling every min vs optimal for two models

**Robust design**

Six points  $\{0,2,4,6,8,10\}$ ,  $w_i=1/8$ , plus  $x=20$ ,  $w_7=1/4$

Efficiency: logistic - 0.79 (was 0.64); exponential - 0.87 (was 0.71)

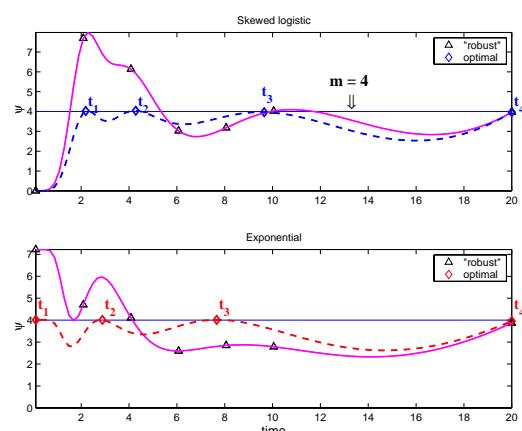


Figure 5: Robust design for two models

### Optimal Design with Cost Constraints

1. Standard normalization:  $N$  - available resource,

$$\sum_i n_i = N \implies M_\xi(\theta) = \sum_{i=1}^n \frac{n_i}{N} \mu(x_i, \theta).$$

2. Let each measurement at point  $x_i$  be associated with cost  $c(x_i)$ ,

$$\sum_i n_i c(x_i) \leq C.$$

Now normalize the information by the total cost  $C$ :

$$M_C(\theta) = \sum_{i=1}^n \frac{n_i}{C} \mu(x_i, \theta) = \sum_i w_i \tilde{\mu}(x_i, \theta),$$

$w_i = n_i c(x_i)/C$ ;  $\tilde{\mu}(x_i, \theta) = \mu(x_i, \theta)/c(x_i) \implies$  same framework.

From weights  $w_i$  to frequencies:  $n_i = [w_i C / c(x_i)]$ .

### Motivating Example 3 (cont. from pp.5-7)

$$\eta(x, \gamma) = f_1(x)/V_1 = \frac{K_a D}{V_1} [b_0 e^{-\alpha_0 x} + b_1 e^{-\alpha_1 x} + b_2 e^{-\alpha_2 x}],$$

$D$  - dose; coefficients  $b_i, \alpha_i$  are expressed via rate constants  $K_a, K_{10}, K_{12}, K_{21}$ .

Seber and Wild (1989)

- $\underline{\gamma} = (V_1, K_a, K_{10}, K_{12}, K_{21})^T$  - response parameters,  $\dim(\gamma) = m_\gamma$ .

- $\gamma_j$  - parameters of individual  $j$  (sampled from population)

$$E\gamma_j = \gamma^0, \quad Var(\gamma_j) = \Lambda$$

$\Delta$  - variance parameters,  $(m_\gamma \times m_\gamma)$  matrix.

- Data  $y(x_{ij}) = \eta(x_{ij}, \gamma_j) + \varepsilon_{ij}$ ,  $i = 1, \dots, k_j$  ( $k_j = 1$  or 2).

Errors  $\varepsilon_{ij}$ : i.i.d. with zero mean and variance  $\sigma^2$ .

$\sigma^2$  - error parameter.

All model parameters:

$$\theta = (\gamma; \Lambda; \sigma^2)$$

If  $X_j = (x_{1,j}, \dots, x_{k_j,j})$  - time points for individual  $j$ , then

$$\begin{aligned} S(X_j, \theta) &= Var(X_j, \theta) = F^T(X_j, \gamma) \Lambda F(X_j, \gamma) + \sigma^2 I_{k_j}, \\ F(X_j, \gamma) &= \left[ \frac{\partial \eta(x_{ij}, \gamma)}{\partial \gamma} \right] - (m_\gamma \times k_j) - \text{matrix} \end{aligned}$$

Now use general formulas (p.9). To implement cost constraints, take

$$\tilde{\mu}(X_j, \theta) = \mu(X_j, \theta)/C_1, \quad \text{if } k_j = 1,$$

$$\tilde{\mu}(X_j, \theta) = \mu(X_j, \theta)/C_2, \quad \text{if } k_j = 2.$$

Details: Fedorov, Gagnon, and Leonov (2001)

Example (cont). Dose administered at  $x = 0$ , sampling allowed every hour from  $x = 1$  to  $x = 20$ .

If  $2^{nd}$  sample taken at  $x_2$ , then samples are separated by at least some positive interval:

$$x_2 - x_1 \geq \Delta \quad (\text{here } \Delta = 4) \implies \mathbf{X} = Z_1 \cup Z_2,$$

Singles:  $Z_1 = \{ x \in [1, 20] \}$ ,

Pairs:  $Z_2 = \{ (x_1, x_2) : x_2 \geq x_1 + 4; x_1, x_2 \in [1, 20] \}$ .

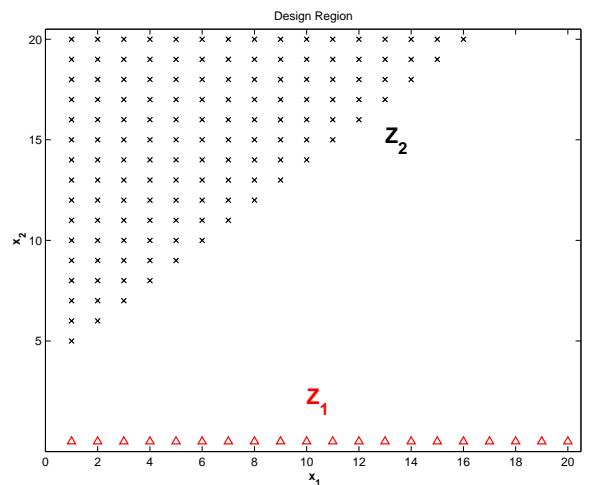


Figure 6: Design Region: singles  $Z_1$  and pairs  $Z_2$

For this example:

$$\theta = (\gamma_1, \gamma_2; \gamma_3, \gamma_4; \lambda_1, \lambda_2, \lambda_3, \lambda_4; \sigma^2)^T, \quad \dim(\theta) = 9.$$

For locally optimal designs, parameter estimates obtained from data.

Example 1.1: singles only  $\implies$

$$\{w_1 = 0.22, w_3 = 0.2, w_7 = 0.19, w_{13} = 0.18, w_{20} = 0.21\}.$$

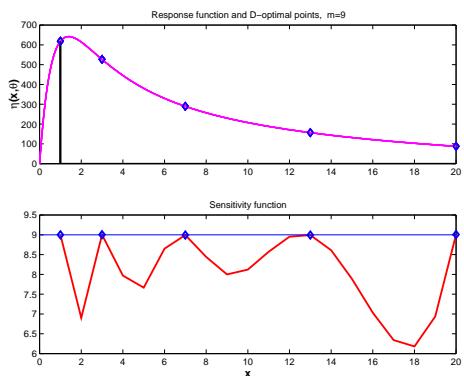


Figure 7: Response and sensitivity functions, single samples ONLY

Example 1.2: equal cost for single samples and pairs,  $C_1 = C_2$ .

D-optimal design: three pairs,

$$X_1 = (1, 5), w_1 = 0.39; X_2 = (3, 11), w_2 = 0.27; X_3 = (10, 20), w_3 = 0.34.$$

Sensitivity function for subregion  $Z_1$  is well below  $m = 9$  (not printed)

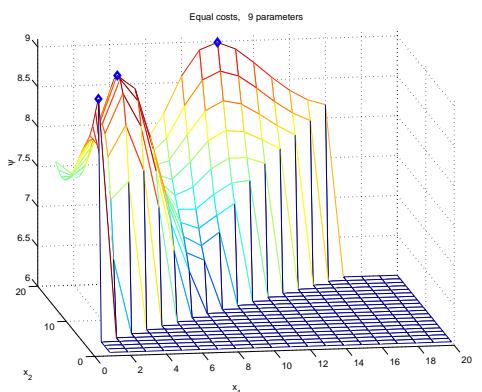


Figure 8: Sensitivity function for subregion  $Z_2$  (pairs)

Example 1.3:  $C_1=1$ ,  $C_2=10$  - pairs cost more.

D-optimal design is supported on six points:

**Four** single points  $(1,3,8,20)$ , weights  $w_{1-4}=(0.2, 0.21, 0.2, 0.19)$ ;

**Two** pairs:  $X_5=(1,6)$ ,  $w_5=0.09$ ;  $X_6=(14,18)$ ,  $w_6=0.11$ .

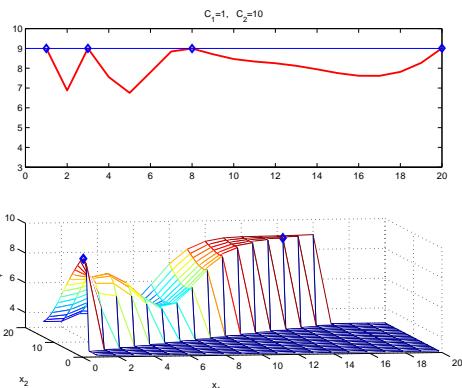


Figure 9: Sensitivity function for subregions  $Z_1$  (top) and  $Z_2$  (bottom)

Current project: PK study, bolus infusion at  $x = 0 \implies$

$$\begin{cases} \dot{f}_1(t) = -(K_{12} + K_{10}) f_1(t) + K_{21} f_2(t) \\ \dot{f}_2(t) = K_{12} f_1(t) - K_{21} f_2(t) \\ f(0) = (D, 0) \end{cases}$$

Time:  $[0, 144]$  hours, current design - 16 samples over the course of 6 days:

$$\{5, 15, 30, 45\} \text{ min}; \{1, 2, \dots, 144\} \text{ hours}$$

Number of candidate “sets” is large: binomial coefficients

$$C(16, 8) = 12870; C(16, 9) = C(16, 7) = 11440$$

Constraint: amount of blood drawn  $\implies$  cost function.

## REFERENCES

35

### References

- [1] Atkinson, A.C. and Cook, R.D. (1995), D-optimum designs for heteroscedastic linear models, *JASA*, **90** (429), 204-212.
- [2] Atkinson, A.C., and Donev, A. (1992), *Optimum Experimental Designs*, Clarendon Press, Oxford.
- [3] Beal, S.L., Sheiner, L.B. (1988), Heteroscedastic nonlinear regression. *Technometrics*, **30**(3), 327-338.
- [4] Bezeau, M., Endrenyi, L. (1986), Design of experiments for the precise estimation of dose-response parameters: the Hill equation. *Journal of Theoretical Biology*, **123**, 415-430.
- [5] Davidian, M., and Carroll, R.J. (1987), Variance function estimation, *JASA*, **82** (400), 1079-1091.
- [6] Downing, D.J., Fedorov, V.V., Leonov, S. (2001), Extracting information from the variance function: optimal design. In: Atkinson, A.C., Hackl, P., Müller, W.G. (eds.) *MODA6 - Advances in Model-Oriented Design and Analysis*, Heidelberg, Physica-Verlag, 45-52.
- [7] Fedorov, V.V. (1974), Regression problems with controllable variables subject to error. *Biometrika*, **61**, 49-56.
- [8] Fedorov, V., Gagnon, R., Leonov, S. (2001), *Optimal Design for Multiple Responses with Variance Depending on Unknown Parameters*, GlaxoSmithKline BDS Technical Report 2001-03.
- [9] Fedorov, V.V., and Hackl, P. (1997), *Model-Oriented Design of Experiments*, Springer-Verlag, New York.
- [10] Ford, I., Torsney, B., and Wu, C.F.J. (1992), The use of a canonical form in the construction of locally optimal designs for non-linear problems. *J. Royal Stat.Soc., Ser. B*, **54** (2), 569-583.
- [11] Hedayat, A.S., Yan, B., Pezutto, J.M. (1997), Modeling and identifying optimum designs for fitting dose response curves based on raw optical data. *J. Amer. Stat. Assoc.*, **92**, 1132-1140.

## REFERENCES

36

- [12] Jennrich, R.I. (1969), Asymptotic properties of nonlinear least squares estimators, *Ann. Math. Stat.* **40**, 633-643.
- [13] Jennrich, R.I., and Schluchter, M.D. (1986), Unbalanced repeated-measures models with structured covariance matrices, *Biometrics*, **42**, 805-820.
- [14] Karpinski, K.F. (1990), Optimality assessment in the enzyme-linked immunosorbent assay (ELISA). *Biometrics*, **46**, 381-390.
- [15] Kiefer, J., and Wolfowitz, J. (1960), The equivalence of two extremum problems, *Canadian Journal of Mathematics*, **12**, 363-366.
- [16] Magnus, J.R., and Neudecker, H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley, New York.
- [17] Malyutov, M.B. (1988), Design and analysis in generalized regression model F, In: Fedorov, V.V., Läuter, H. (Eds.), *Model-Oriented Data Analysis*, Springer-Verlag, Berlin, 72-76.
- [18] Mentré, F., Mallet, A., Baccar, D. (1997), Optimal design in random-effects regression models, *Biometrika*, **84** (2), 429-442.
- [19] Seber, G.A.F., Wild, C.I. (1989), *Nonlinear Regression*, Wiley, New York.
- [20] Vonesh, E.F., Chinchilli, V.M. (1997), *Linear and Nonlinear Models for the Analysis of Repeated Measurements*, Marcel Dekker, New York.
- [21] White, L.V. (1975), *The optimal design of experiments for estimation in nonlinear model*. Ph. D. Thesis, University of London.
- [22] Wu, C.F.J. (1981), Asymptotic theory of nonlinear least squares estimation, *Ann. Stat.*, **9** (3), 501-513.
- [23] Wu, C.F.J. (1988), Optimal design for percentile estimation of a quantal-response curve. In: Dodge, Y., Fedorov, V.V., Wynn, H.P. (eds.), *Optimal Design and Analysis of Experiments*. Elsevier, North Holland, 213-223.